A Decentralized Framework for the Optimal Coordination of Distributed Energy Resources

A Decentralized Framework for the Optimal Coordination of Distributed Energy Resources

Miguel F. Anjos, *Senior Member, IEEE,* Andrea Lodi, and Mathieu Tanneau

***Abstract*—Demand-response aggregators are faced with the challenge of how to best manage numerous and heterogeneous Distributed Energy Resources (DERs). This paper proposes a decentralized methodology for optimal coordination of DERs. The proposed approach is based on Dantzig-Wolfe decomposition and column generation, thus allowing to integrate any type of resource whose operation can be formulated within a mixed-integer linear program. We show that the proposed framework offers the same guarantees of optimality as a centralized formulation, with the added benefits of distributed computation, enhanced privacy, and higher robustness to changes in the problem data. The practical efficiency of the algorithm is demonstrated through extensive computational experiments, on a set of instances generated using data from Ontario energy markets. The proposed approach was able to solve all test instances to proven optimality, while achieving significant speed-ups over a centralized formulation solved by state-of-the-art optimization software.**

***Index Terms*—Column generation, Dantzig-Wolfe decomposition, demand response aggregation, distributed energy resources, mixed-integer linear programming, smart grid**

NOMENCLATURE

Parameters

C_d^{th}	Thermal capacity of thermal load d.
$E_d^{\text{bat, min}}$	Minimum state of charge of battery device d.
$E_d^{\text{bat, max}}$	Maximum state of charge of battery device d.
L_d^{uni}	Cycle duration of uninterruptible device d.
$\tilde{P}_{d,t}^{\text{unc}}$	Power consumption of uncontrollable load d during time period t.
$\tilde{P}_{d,t}^{\text{cur}}$	Power consumption of curtailable load d during time period t, in the absence of curtailment.
$\tilde{P}_{d,l}^{\text{uni}}$	Power requirement of uninterruptible load d, during its cycle's phase l.
$P_d^{\text{th, min}}$	Minimum power consumption of thermal load d, when that device is on.
$P_d^{\text{th, max}}$	Maximum power consumption of thermal load d, when that device is on.
$P_d^{\text{ch, min}}$	Minimum charging power of battery device d.
$P_d^{\text{ch, max}}$	Maximum charging power of battery device d.
$P_d^{\text{dis, min}}$	Minimum discharging power of battery device d.
$P_d^{\text{dis, max}}$	Maximum discharging power of battery device d.
$P_r^{\min}$	Minimum value of household r's net load.
$P_r^{\max}$	Maximum value of household r's net load.
$P_a^{\min}$	Minimum value of the aggregated residential net load.
$P_a^{\max}$	Maximum value of the aggregated residential net load.
R	Number of resources.
T	Length of the time-horizon.
$\Delta\tau$	Duration of each time period.
η_d^{ch}	Charging efficiency of battery device d, in $[0, 1)$.
η_d^{dis}	Discharging efficiency of battery device d, in $[0, 1)$.
η_d^{th}	Thermal efficiency of thermal load d.
$\Theta_d^{\min}$	Minimum temperature requirement of thermal load d.
$\Theta_d^{\max}$	Maximum temperature requirement of thermal load d.
$\Theta_{d,t}^{\text{ext}}$	Outside temperature for thermal load d during time period t.
μ_d^{th}	Thermal conductivity of thermal load d.
Π_t	Market price of electricity during time period t.

Sets

$\mathcal{D}_r^{unc}$	Set of uncontrollable loads of household r.
$\mathcal{D}_r^{cur}$	Set of curtailable loads of household r.
$\mathcal{D}_r^{uni}$	Set of uninterruptible loads of household r.
$\mathcal{D}_r^{def}$	Set of deferrable loads of household r.
$\mathcal{D}_r^{th}$	Set of thermal loads of household r.
$\mathcal{D}_r^{sto}$	Set of energy storage devices of household r.
$\mathcal{R}$	Set of resources $\mathcal{R} = \{1, ..., R\}$.
$\mathcal{T}$	Time horizon $\mathcal{T} = \{0, ..., T-1\}$.
$\mathcal{X}_r$	Set of feasible operational schedules of resource r.
Γ_r	Set of extreme rays of $conv(\mathcal{X}_r)$.
Ω_r	Set of extreme vertices of $conv(\mathcal{X}_r)$.

Variables

$e_{d,t}^{\text{bat}}$	State of charge of battery device d at the end of time period t.
$p_{a,t}$	Net aggregated load during time period t.
$p_{d,t}$	Algebraic power consumption of device d during time period t.
$p_{d,t}^{\text{ch}}$	Charging power of battery device d during time period t.
$p_{d,t}^{\text{dis}}$	Discharging power of battery device d during time period t.
$p_{r,t}$	Net load of household r during time period t.
$u_{d,t}^{\text{ch}}$	Binary variable that takes value 1 (resp. 0) if battery device d is charging (resp. not charging) during time period t.

The authors are with École Polytechnique de Montréal, Department of Mathematics and Industrial Engineering, Montréal, QC H3T 1J4, Canada, & Groupe d'études et de recherche en analyse des décisions (GERAD), Montréal, QC H3T 1J4, Canada (e-mail: anjos@stanfordalumni.org).

$u^{\text{dis}}_{d,t}$	Binary variable that takes value 1 (resp. 0) if battery device d is discharging (resp. not discharging) during time period t.
$u^{\text{cur}}_{d,t}$	Binary variable that takes value 1 (resp. 0) if curtailable load d is on (resp. off) during time period t.
$u^{\text{th}}_{d,t}$	Binary variable that takes value 1 (resp. 0) if thermal load d is on (resp. off) during time period t.
$v^{\text{uni}}_{d,t}$	Binary variable that takes value 1 if uninterruptible load d's cycle is started at time t, and 0 otherwise.
$\mathbf{x}_r$	Vector of decision variables (operational schedule) of resource r.
$\mathbf{y}$	Vector of decision variables for the aggregator.
$\theta^{\text{th}}_{d,t}$	Temperature of thermal load d during time period t.
λ	Vector of primal variables in the Master Problem.
π	Vector of dual variables (shadow prices).
σ	Vector of dual variables (shadow marginal costs).

I. Introduction

SMART grids hold the promise of more reliable and sustainable power grids, through the integration of communication technologies, advanced computation and intelligent controls. Among smart grid-enabled paradigms, Demand Response (DR) programs [1] enable end-users to dynamically adjust their electricity consumption, in response to price signals or incentives. The present work focuses on the DR potential of Distributed Energy Resources (DERs) that include, among others, flexible loads, distributed generation, and distributed energy storage [2].

Because of their small size, individual resources have a negligible marginal impact at the grid level. This has motivated the introduction of aggregators [3], that act as intermediaries between the grid and resources, thus enabling the latters to participate in traditional energy markets. Therefore, aggregators must address the challenges associated with coordinating numerous and heterogeneous resources, whose operation may involve discrete decisions. To that end, various coordination mechanisms have been investigated in the literature [4]–[6], and essentially break down into two main categories: price-based and incentive-based mechanisms.

In price-based mechanisms, price signals are communicated to users, either periodically, e.g. in Time-of-Use rates, or dynamically as in the case of Real-Time Pricing [6]. Users then adapt their consumption by optimizing their own individual objective, hopefully reducing electricity consumption when prices are higher. Therefore, the efficiency of a price-based DR mechanism is contingent on how price signals are computed and how users react to them. In [7]–[10], the DR problem is formulated as a Stackelberg game between an electricity retailer (leader) and electricity consumers (followers). This results in a bi-level problem wherein the retailer sets prices so as to maximize its own payoff, while accounting for consumers' optimal response to those prices. While authors in [8] and [10] focus on Time-Of-Use rates, Dynamic pricing is also considered in [7] and [9]. These works perform a reformulation of the bi-level program as a single-level Mixed-Integer Linear Program, which can then be solved using standard optimization software. Nevertheless, this approach suffers from poor scalability, and does not handle the presence of discrete variables in the consumers' operation. Another methodology, employed in [11]–[14], formulates the DR problem as a social welfare maximization problem. In [11] and [12], the total cost of electricity over a finite time-horizon is minimized, while authors in [13] and [14] consider individual utility functions for electricity consumers. All of [11]–[14] reduce to solving a convex (continuous) optimization problem, where dual information (i.e., Lagrange multipliers) is used to compute the price signals. However, this methodology does not apply when discrete variables are present, which is the setting we consider in this work.

On the other hand, incentive-based mechanisms reward users who are willing to participate in DR programs, for example by offering discounts on electricity bills [1], or on a per-event basis [15]. Incentive-based mechanisms include Direct Load Control schemes, which is the approach considered in this paper. Specifically, we focus on jointly coordinating the operation of a (large) number of DERs, so as to optimize a global objective. This can be formulated as a centralized optimization problem, as was explored in [16]–[19], wherein a central controller manages each individual resource. Although a centralized formulation offers the strongest optimality guarantees, it quickly becomes intractable when the number of resources increases, both due to memory requirements and the presence of discrete variables. Furthermore, the disclosure of private information by the resources raises privacy concerns, and puts additional burden on communication requirements.

Decentralized methods, on the other hand, distribute the computational effort among resources by leveraging local computing power. This typically results in better scalability, reduced communication overheads, and improved privacy. Several distributed heuristics were proposed, for example in [20]–[23], but provide weaker guarantees of optimality, since heuristic methods converge at best to a local optimum. Nevertheless, classical decomposition techniques allow to solve the centralized problem in a distributed way, thus offering the same guarantees of optimality, with the added benefit of decentralized computation. Therefore, in this work, we focus on exact, decomposition-based methods.

A large body of literature has focused on dual decomposition and Lagrangian-based methods. Dual decomposition yields a separable structure, which in turn enables distributed implementations. The related works [24]–[29] thus differ mainly in which algorithm is used to optimize the dual Lagrangian. In [24] and [25], the authors consider an augmented Lagrangian-based relaxation, which is formulated as a consensus problem and solved with the Alternating Direction Method of Multipliers (ADMM). However, ADMM does not handle discrete variables, that are used to model on-off constraints. In a similar fashion, standard Lagrangian relaxation is used in [26]–[29]. A classical sub-gradient algorithm is investigated in [26] and cutting-planes methods are studied in [27], but discrete variables were not considered. Although the works in

[28] and [29] do consider mixed-integer variables, they rely on recovery heuristics to obtain feasible solutions. A bundle method is used in [28], while a double smoothing of the dual Lagrangian is applied in [29], allowing the use of a more efficient gradient-based algorithm. Overall, the main drawback of Lagrangian-based approaches is the recovery of feasible solutions when strong duality does not hold, which is generally the case for mixed-integer problems.

Alternatively, Dantzig-Wolfe (DW) decomposition-based approaches were investigated in [30]–[33]. In [30] and [31], DW decomposition is applied to demand-response problems related to peak-load management and, in [32], to the charging of electric vehicles. However, only a few types of devices were considered in these works, with no discrete variables involved. Finally, a column generation-based heuristic is used in [33] to schedule residential heating systems. Nevertheless, this heuristic approach does not provide a guarantee of optimality.

A. Contribution and outline

In this paper, we propose a scalable and exact methodology for the coordination of DERs whose operation involves discrete decisions, i.e., mixed-integer variables. Our proposed framework is based on DW decomposition and Column Generation, and has several advantageous features compared to existing approaches, which are either heuristic or poorly scalable. First, DW decomposition provides a systematic and efficient treatment of discrete variables. This yields a technology-agnostic formulation as shown in Section III-A. Second, our column-generation algorithm naturally distributes among resources, which makes it highly scalable, as demonstrated in Section IV. Third, this decentralization offers enhanced privacy to the resources. Fourth, our formulation is robust to numerical changes in the problem data, with empirical evidence reported in Section IV-E.

The rest of the paper is organized as follows. In Section II, we introduce operational models, and formulate the aggregation problem as a (centralized) Mixed-Integer Linear Program. We then present our methodology in Section III, including the Dantzig-Wolfe reformulation in III-A and our column-generation algorithm in III-B and III-C. Computational results are reported in Section IV, and Section V concludes the paper.

Unless specified otherwise, durations are expressed in hours (h), power and energy quantities in kilowatt (kW) and kilowatt-hour (kWh) respectively, and temperatures in degrees Celsius. Energy prices are given in dollars per kilowatt-hour ($/kWh).

II. Problem formulation

We consider a set of households that interact with an aggregator, in order to minimize the total cost of purchasing electricity from the grid. The aggregator and households are assumed to behave rationally, and households do not interact directly with each other. Finally, the environment is assumed to be deterministic. This last assumption is further discussed in Section IV.

Operational models are presented in II-A and II-B, and the aggregation problem is formulated in II-C.

A. Device constraints

Following the classification in [34], devices are grouped into six classes: uncontrollable loads (unc), curtailable loads (cur), uninterruptible loads (uni), deferrable loads (def), thermal loads (th), and energy storage (bat). The sequence $\mathbf{p}_d = (p_{d,0}, \cdots, p_{d,T-1})$ is called the device's load profile. A negative consumption, i.e., $p_{d,t} < 0$, indicates the device generates power.

Uncontrollable loads include devices whose operation cannot be altered. The load profile of such a device d is thus

$$p_{d,t} = \tilde{P}^{\text{unc}}_{d,t}. \qquad \forall t \in \mathcal{T} \tag{1}$$

Curtailable loads refer to devices whose operation at a given time t can be altered, independently of past and future operations. The operation of a curtailable load d is

$$p_{d,t} - u^{\text{cur}}_{d,t}\tilde{P}^{\text{cur}}_{d,t} = 0, \qquad \forall t \in \mathcal{T} \tag{2}$$

$$u^{\text{cur}}_{d,t} \in \{0,1\}. \qquad \forall t \in \mathcal{T} \tag{3}$$

On-off curtailment is modelled through binary variables $u^{\text{cur}}_{d,t}$. Continuous curtailment is modelled by relaxing the integrality constraint on $u^{\text{cur}}_{d,t}$. This model can be extended to include several operation modes. Finally, a penalty is typically incurred when demand is curtailed. However, we do not explicitly formulate it here because, as will be detailed in Section IV, we do not consider in our experiments any demand that can be curtailed.

Uninterruptible loads are comprised of devices such as dishwashers and clothes dryers, whose operation is typically composed of a finite number of cycles. A cycle's start-up time is flexible but, once started, it cannot be interrupted. For simplicity, we consider a single cycle. The device's operation is then defined by

$$p_{d,t} - \sum_{l=0}^{L-1} v^{\text{uni}}_{d,t-l}\tilde{P}^{\text{uni}}_{d,l} = 0, \qquad \forall t \in \mathcal{T} \tag{4}$$

$$\sum_{t=0}^{T-L} v^{\text{uni}}_{d,t} = 1, \tag{5}$$

$$v^{\text{uni}}_{d,t} \in \{0,1\}, \qquad \forall t \in \mathcal{T} \tag{6}$$

where constraint (5) ensures the cycle is started exactly once, and $v^{\text{uni}}_{d,t} = 0, \forall t < 0$. This framework naturally extends to several cycles with precedence constraints as proposed in [35].

Deferrable loads include devices such as an electric vehicle's (EV) charger, whose consumption may be shifted earlier or later in time. Deferrable loads are modelled as follows:

$$E^{\text{def, min}}_d \leq \Delta\tau \sum_{t\in\mathcal{T}} p_{d,t} \leq E^{\text{def, max}}_d, \tag{7}$$

$$u^{\text{def}}_{d,t} P^{\text{def, min}}_d \leq p_{d,t} \leq u^{\text{def}}_{d,t} P^{\text{def, max}}_d, \qquad \forall t \in \mathcal{T} \tag{8}$$

$$u^{\text{def}}_{d,t} \in \{0,1\}. \qquad \forall t \in \mathcal{T} \tag{9}$$

On-off constraints are modelled by (8)-(9).

Thermal loads encompass devices like space heaters and air conditioners, that aim at keeping a system's (e.g., a room)

temperature within a certain range. Their operation is modelled by

$$\Theta_d^{\min} \leq \theta_{d,t}^{\text{th}} \leq \Theta_d^{\max}, \qquad \forall t \in \mathcal{T} \quad (10)$$

$$\frac{\mu_d^{\text{th}}}{C_d^{\text{th}}}(\Theta_{d,t}^{\text{ext}} - \theta_{d,t}^{\text{th}}) + \frac{\eta_d^{\text{th}}}{C_d^{\text{th}}} p_{d,t} = \frac{\theta_{d,t+1}^{\text{th}} - \theta_{d,t}^{\text{th}}}{\Delta\tau}, \quad \forall t \in \mathcal{T} \quad (11)$$

$$u_{d,t}^{\text{th}} P_d^{\text{th, min}} \leq p_{d,t} \leq u_{d,t}^{\text{th}} P_d^{\text{th, max}}, \qquad \forall t \in \mathcal{T} \quad (12)$$

$$u_{d,t}^{\text{th}} \in \{0,1\}. \qquad \forall t \in \mathcal{T} \quad (13)$$

The evolution of the system's temperature in (11) is given by the first-order approximation of a thermodynamic model [36]. The device's on-off constraints are modelled in (12)-(13).

Energy storage devices can store energy and release it later. For simplicity, we consider the case of batteries, for which an operational model (adapted from [37]) is

$$E_d^{\text{bat, min}} \leq e_{d,t}^{\text{bat}} \leq E_d^{\text{bat, max}}, \qquad \forall t \in \mathcal{T} \quad (14)$$

$$p_{d,t} - p_{d,t}^{\text{ch}} + p_{d,t}^{\text{dis}} = 0, \qquad \forall t \in \mathcal{T} \quad (15)$$

$$\Delta\tau \left(\eta_d^{\text{ch}} p_{d,t}^{\text{ch}} - \frac{1}{\eta_d^{\text{dis}}} p_{d,t}^{\text{dis}} \right) = e_{d,t}^{\text{bat}} - e_{d,t-1}^{\text{bat}}, \quad \forall t \in \mathcal{T} \quad (16)$$

$$u_{d,t}^{\text{ch}} P_d^{\text{ch, min}} \leq p_{d,t}^{\text{ch}} \leq u_{d,t}^{\text{ch}} P_d^{\text{ch, max}}, \qquad \forall t \in \mathcal{T} \quad (17)$$

$$u_t^{\text{dis}} P_d^{\text{dis, min}} \leq p_{d,t}^{\text{dis}} \leq u_{d,t}^{\text{dis}} P_d^{\text{dis, max}}, \qquad \forall t \in \mathcal{T} \quad (18)$$

$$u_{d,t}^{\text{ch}} + u_{d,t}^{\text{dis}} \leq 1, \qquad \forall t \in \mathcal{T} \quad (19)$$

$$u_{d,t}^{\text{ch}}, u_{d,t}^{\text{dis}} \in \{0,1\}. \qquad \forall t \in \mathcal{T} \quad (20)$$

The internal dynamics of the battery are captured by (16), and on-off constraints (17)-(20) ensure that the battery cannot be simultaneously charged and discharged. This model can be further extended to include constraints on ramping and cycling, see, e.g. [37].

Finally, renewable generation can be modelled as a negative load. Depending on systems' specifications, it may be either uncontrollable or curtailable.

B. Household and aggregator constraints

We now consider a given household $r \in \mathcal{R}$. Additional constraints at the household level are formulated as

$$P_r^{\min} \leq p_{r,t} \leq P_r^{\max}, \qquad \forall t \in \mathcal{T} \quad (21)$$

$$p_{r,t} = \sum_{d \in \mathcal{D}_r} p_{d,t}. \qquad \forall t \in \mathcal{T} \quad (22)$$

corresponding to that house's circuit breaker's operating range.

Similarly, the aggregator must ensure that the aggregated load $p_{a,t}$ remains within physical limitations, i.e.,

$$P_a^{\min} \leq p_{a,t} \leq P_a^{\max}, \qquad \forall t \in \mathcal{T} \quad (23)$$

$$p_{a,t} = \sum_{r \in \mathcal{R}} p_{r,t}. \qquad \forall t \in \mathcal{T} \quad (24)$$

If households were operated independently, i.e., without interacting with the aggregator, constraints (23) may be violated. In practice, this could result in equipment damage, or localised blackouts.

C. Aggregation problem

The aggregation problem consists here in minimizing the total cost of purchasing energy from the grid, while satisfying all operational constraints. It is formulated as the following Mixed-Integer Linear Program (MILP):

$$\begin{aligned} \min \quad & \sum_{t \in \mathcal{T}} \Pi_t \times \Delta\tau \times p_{a,t} && (25) \\ \text{s.t.} \quad & (1), && \forall r, \forall d \in \mathcal{D}_r^{unc} \\ & (2)-(3), && \forall r, \forall d \in \mathcal{D}_r^{cur} \\ & (4)-(6), && \forall r, \forall d \in \mathcal{D}_r^{uni} \\ & (7)-(9), && \forall r, \forall d \in \mathcal{D}_r^{def} \\ & (10)-(13), && \forall r, \forall d \in \mathcal{D}_r^{th} \\ & (14)-(20), && \forall r, \forall d \in \mathcal{D}_r^{sto} \\ & (21)-(22), && \forall r \\ & (23)-(24). \end{aligned}$$

We now introduce a more general and compact notation for the aggregation problem (25). For each household r, let $\mathbf{x}_r$ denote its vector of decision variables, i.e., the vector obtained by concatenating all decision variables specific to that household. Here, $\mathbf{x}_r$ would be composed of a household's net load, plus the decision variables of each device in that household. In what follows, we refer to $\mathbf{x}_r$ as the operational schedule of household r. Define $\mathcal{X}_r$ as the set of all feasible operational schedules for household r, so that operational constraints for that household are written in the compact form $\mathbf{x}_r \in \mathcal{X}_r$.. Therefore, $\mathcal{X}_r$ is defined by a finite number of linear constraints and integrality requirements, so that $conv(\mathcal{X}_r)$ is a polyhedron. Finally, a household's operating cost is written $c_r^T x_r$ for a given cost vector c_r. In the present case, we have $c_r = 0$ for every r, since no household-specific cost is considered. Nevertheless, we keep the objective term c_r in the formulation for generalization purposes.

Similarly, let $\mathbf{y}$ denote the vector of decision variables that are specific to the aggregator. For the case at hand, $\mathbf{y}$ corresponds to the aggregated load profile $\mathbf{p}_a$. The aggregator's operating cost is written $q^T\mathbf{y}$, while constraints (23)-(24) are written $M\mathbf{y} + \sum_r A_r \mathbf{x}_r = b$, without loss of generality.

The aggregation problem can therefore be written in the general compact form

$$\min_{\mathbf{y}, \mathbf{x}_1, \ldots, \mathbf{x}_R} \quad q^T\mathbf{y} + \sum_{r \in \mathcal{R}} c_r^T \mathbf{x}_r \quad (26)$$

$$\text{s.t.} \quad M\mathbf{y} + \sum_{r \in \mathcal{R}} A_r \mathbf{x}_r = b \quad (27)$$

$$\mathbf{x}_r \in \mathcal{X}_r. \qquad \forall r \quad (28)$$

Constraints (27) induce a coupling between the households, and are thus referred to as linking constraints. Conversely, constraints (28) are separable by household, and are referred to as local constraints. We emphasize that this formulation is not restricted to the aforementioned loads, but allows to integrate any type of DER whose operation can be formulated within a MILP.

Mixed-integer linear programs such as (26)-(28) can be solved to proven optimality using standard optimization soft-

ware. Although MILPs are NP-hard in general, practical instances can often be solved efficiently. Indeed, several of our test instances, with up to hundreds of thousands of variables, were solved to optimality in a few minutes with a state-of-the art solver. Nevertheless, a centralized formulation obviously becomes intractable when dealing with large systems, both due to memory requirements and the presence of discrete variables. Furthermore, as will be demonstrated in Section IV-E, the nature of the DERs may have a significant impact on the formulation's integrality gap, and therefore on computing time.

III. Distributed Column-Generation framework

We now present a decentralized framework for solving the aggregation problem to global optimality. To underline the generality of the proposed methodology, we use the notation introduced in Section II-C.

A. Dantzig-Wolfe decomposition

We begin by introducing a Dantzig-Wolfe reformulation of the aggregation problem (26)-(28). First, for each resource $r \in \mathcal{R}$, let Ω_r (resp. Γ_r) denote the set of extreme vertices (resp. extreme rays) of $conv(\mathcal{X}_r)$. Both sets are well-defined and finite since, as mentioned in Section II-C, $conv(\mathcal{X}_r)$ is a polyhedron. For simplicity, we assume that $\mathcal{X}_r$ is bounded, so that $\Gamma_r = \emptyset$.

We then apply the Dantzig-Wolfe decomposition principle [38] to the aggregation problem (26)-(28), which yields the extended formulation

$$\min_{\mathbf{y},\lambda} \quad q^T\mathbf{y} + \sum_{r\in\mathcal{R},\omega\in\Omega_r} c_{r,\omega}\lambda_{r,\omega} \tag{29}$$

$$\text{s.t.} \quad M\mathbf{y} + \sum_{r\in\mathcal{R},\omega\in\Omega_r} \lambda_{r,\omega}a_{r,\omega} = b, \tag{30}$$

$$\sum_{\omega\in\Omega_r} \lambda_{r,\omega} = 1, \qquad \forall r \tag{31}$$

$$\lambda \geq 0, \tag{32}$$

$$\sum_{\omega\in\Omega_r} \lambda_{r,\omega}\omega = \mathbf{x}_r, \qquad \forall r \tag{33}$$

$$\mathbf{x}_r \in \mathcal{X}_r, \qquad \forall r \tag{34}$$

with the notation $c_{r,\omega} = c_r^T\omega$ and $a_{r,\omega} = A_r\omega$. The objective (29) and linking constraints (30) are simply re-writing of (26) and (27), respectively. Thereby, each extreme point $\omega \in \Omega_r$ is associated to a variable $\lambda_{r,\omega}$, and to a column $a_{r,\omega}$ that corresponds to a load schedule for resource r. The extended formulation (29)-(34) is equivalent to the compact formulation (26)-(28), and is most typically solved using a branch-and-price algorithm. In what follows, we focus on the linear relaxation of the extended formulation, which is given by (29)-(32) and referred to as the Master Problem (MP).

Unlike the centralized formulation (26)-(28), the structure of the MP is independent of the nature of the resources. Indeed, the linking constraints only involve variables corresponding to a resource's net load. In particular, the MP is always a (continuous) linear program, even if a resource's operation involves discrete variables. Therefore, using DW decomposition yields a technology-agnostic formulation. This feature is highly advantageous, since it allows to integrate new types of resources without any major impact on performance.

B. Distributed column generation

Since the MP contains exponentially many variables, it is solved by a Column-Generation (CG) algorithm. We refer to [39] for a thorough overview of column generation and branch-and-price algorithms, as well as the relation between Lagrangian relaxation and Dantzig-Wolfe decomposition.

Consider the Restricted Master Problem (RMP)

$$\min_{\mathbf{y},\lambda} \quad q^T y + \sum_{r\in\mathcal{R},\omega\in\bar{\Omega}_r} c_{r,\omega}\lambda_{r,\omega} \tag{35}$$

$$\text{s.t.} \quad M\mathbf{y} + \sum_{r\in\mathcal{R},\omega\in\bar{\Omega}_r} \lambda_{r,\omega}a_{r,\omega} = b, \tag{36}$$

$$\sum_{\omega\in\bar{\Omega}_r} \lambda_{r,\omega} = 1, \qquad \forall r \tag{37}$$

$$\lambda \geq 0, \tag{38}$$

where, for each resource r, $\bar{\Omega}_r \subset \Omega_r$ denotes the subset of columns that are currently considered. The RMP is initialized with a small number of columns, some of which may be artificial to ensure feasibility.

At the beginning of each iteration, the RMP is solved to optimality. Let π denote the vector of dual variables associated to linking constraints (36), and σ_r the dual variable associated to convexity constraint (37) for resource r. For given r and $\omega \in \Omega_r$, the reduced cost of $\lambda_{r,\omega}$ is then

$$\bar{c}_{r,\omega} = c_r^T\omega - \pi^T A_r\omega - \sigma_r. \tag{39}$$

Therefore, a variable λ_{r,ω^*} with smallest reduced cost is given by the pricing step

$$\omega^* \in \arg\min_{\omega\in\Omega_r} \left(c_r^T\omega - \pi^T A_r\omega - \sigma_r\right). \tag{40}$$

However, explicitly iterating over all Ω_r is prohibitively expensive. Nevertheless, since each $\omega \in \Omega_r$ is an extreme point of conv$(\mathcal{X}_r)$, performing the pricing step (40) is equivalent to solving

$$(SP_r) \quad \omega^* \in \arg\min_{\mathbf{x}_r} \quad (c_r^T - \pi^T A_r)\mathbf{x}_r - \sigma_r \tag{41}$$

$$\text{s.t.} \quad \mathbf{x}_r \in \mathcal{X}_r. \tag{42}$$

The pricing sub-problem (41)-(42) is a small MILP, which we assume can be solved efficiently. If the identified variable λ_{r,ω^*} has negative reduced cost, i.e., $\bar{c}_{r,\omega^*} < 0$, it is added to the RMP. Otherwise, $\bar{c}_{r,\omega^*} \geq 0$ and all variables $\lambda_{r,\omega}, \omega \in \Omega_r$ have non-negative reduced cost. Optimality in the MP is reached when this is the case for all resources, i.e., all variables in the master problem have non-negative reduced cost. The pseudo-code of our CG algorithm is given in Algorithm 1.

Algorithm 1 Column-Generation algorithm

Input: Initial RMP
while stopping criterion not met **do**
  Solve RMP and obtain optimal dual variables (π, σ)
  // Pricing step (distributed)
  for all $r \in \mathcal{R}$ **do**
    Solve SP_r with the query point (π, σ_r), and obtain ω^* such that λ_{r,ω^*} has most negative reduced cost
    if $\bar{c}_{r,\omega^*} < 0$ **then**
      Add corresponding column to the RMP
    end if
  end for
  // Stopping criterion
  if no column added to RMP **then**
    STOP
  end if
end while

All the sub-problems are independent, and therefore can be solved in a decentralized fashion. This offers greater privacy to the resources. Indeed, each sub-problem SP_r is solved locally by resource r. Consequently, the aggregator only has limited information about each resource's operation, in the form of the columns that are generated. Furthermore, this decentralized setting is more robust to communication failure. If a resource r fails to communicate with the aggregator, a locally feasible operation can always be used, e.g. the last column that was generated. Similarly, the aggregator and other resources can carry on the optimization process, considering for resource r only the last column it generated. If communication is later restored, the quality of the final solution will not be affected, although the total number of iterations may obviously increase.

C. Branch-and-price

The MP is the root node of the branching tree in a branch-and-price algorithm. Solving the MP yields a primal optimal solution $(\bar{y}, \bar{\lambda})$ and using it, we define, for each $r \in \mathcal{R}$

$$\bar{\mathbf{x}}_r := \sum_{\omega \in \Omega_r} \bar{\lambda}_{r,\omega} \omega. \tag{43}$$

Since the MP is a relaxation of the extended formulation, the $\bar{\mathbf{x}}_r$ are generally fractional. For example, $\bar{\mathbf{x}}_r$ may represent an operational schedule where a battery is charging and discharging simultaneously, which would violate constraint (20). An optimal integer solution is obtained by branch-and-price.

We implemented a heuristic diving procedure wherein several branching decisions are taken simultaneously. At each successive branching node, a random subset of resources $\tilde{\mathcal{R}}$ is selected. Then, for each $r \in \tilde{\mathcal{R}}$, a feasible $\tilde{\mathbf{x}}_r$ is computed by projecting the current fractional solution $\bar{\mathbf{x}}_r$ onto the feasible set $\mathcal{X}_r$. This is done by solving the MILP

$$\tilde{\mathbf{x}}_r \in \arg\min_{\mathbf{x}_r} \quad \|\mathbf{x}_r - \bar{\mathbf{x}}_r\|_\infty \tag{44}$$

$$\text{s.t.} \quad \mathbf{x}_r \in \mathcal{X}_r. \tag{45}$$

Branching decisions are taken by fixing all integer variables to their value in $\tilde{\mathbf{x}}_r$, and are enforced in the sub-problem SP_r directly. The column corresponding to $\tilde{\mathbf{x}}_r$ is added to the RMP, and all conflicting columns, i.e., columns that do not satisfy the branching decisions, are removed from the formulation. Enforcing branching decisions in the sub-problems results in no conflicting column being generated afterwards. Finally, the relaxation of the newly obtained node is solved by column generation, and the procedure is repeated until all integer variables are fixed. The pseudo-code of this diving procedure is given in Algorithm 2.

Algorithm 2 Diving heuristic procedure

Input: Solution $\bar{\mathbf{x}}$ of the MP, $\bar{\mathcal{R}} := \mathcal{R}$
while stopping criterion not met **do**
  // Branching decisions
  Select a random subset of resources $\tilde{\mathcal{R}} \subset \bar{\mathcal{R}}$
  for all $r \in \tilde{\mathcal{R}}$ **do**
    Compute $\tilde{\mathbf{x}}_r$ and add corresponding column to the RMP
    Remove conflicting columns
    Enforce branching decisions in SP_r
  end for
  Solve node relaxation using Algorithm 1, obtain new $\bar{\mathbf{x}}$
  $\bar{\mathcal{R}} := \bar{\mathcal{R}} - \tilde{\mathcal{R}}$
  // Stopping criterion
  if $\bar{\mathcal{R}} = \emptyset$ **then**
    STOP
  end if
end while

This heuristic was run right after solving the MP and, in all test instances, it was able to find what turned out to be an optimal integer solution. More precisely, a feasible integer solution was always found by the heuristic, and its optimality was proven using the lower bound provided by the MP. Therefore, no branching was needed, which we emphasize may not be true in general.

IV. Computational results

We now report computational results for the proposed column-generation algorithm (CG), using the centralized formulation (MILP) as a baseline. Tests are carried out on a set of 1120 instances, generated using data from Ontario energy markets. We first describe our test methodology in IV-A and IV-B. Numerical results are analysed in IV-C and IV-D, and the robustness of the formulation is further assessed in IV-E.

For reproducibility, we make our code for generating instances and numerical data publicly available[1].

A. Numerical instantiation

For all simulations, the time-horizon begins at 5am Monday, January 18^{th} 2016, and one time-period is always $\Delta\tau = 1$h. We consider four different values for T and seven for R, which are indicated in Figure 1 and Table II, respectively. Ontario's provincial load, production and pricing data are obtained from [40].

The set of considered devices and their ownership rates are given in Table I (adapted from [34]). The ownership rate of a device is interpreted as the probability that that device be present in a given household. Devices that are not explicitly considered in this work are aggregated into one uncontrollable load for each household. In addition, we consider correlation

[1] https://github.com/mtanneau/DER_instances

Table I
DEVICES CLASSIFICATION AND OWNERSHIP RATES

Device	Classification	Own. rate (%)
Dishwasher	Uninterruptible load	65
Clothes washer	Uninterruptible load	90
Clothes dryer	Uninterruptible load	75
Electric heating	Thermal load	60
Electric vehicle	Deferrable load	ξ^{dep}
Home battery	Energy storage	ξ^{dep}
Rooftop solar	Curtailable load	ξ^{dep}
Others	Uncontrollable load	100

among devices. It is assumed that only households with a clothes washer own a clothes dryer. Because available data regarding the penetration of EVs, batteries and rooftop solar in Ontario is very scarce, we assumed that households own either none or the three of them. Several deployment scenarios are considered, corresponding to ownership rates $\xi^{\text{dep}} = 0\%, 33\%, 66\%$ and 100%.

Dishwashers, clothes washers and clothes dryers are modelled as uninterruptible loads. Each appliance's operation consists of one cycle per day, with a constant power consumption of 1kW. Cycles' durations are 2h for dishwashers, 2h for clothes washers, and 3h for clothes dryers.

For electric heating systems, the outside temperature Θ^{ext} is obtained by perturbing a reference profile Θ^{ref} that consists of hourly readings from a weather station in the Toronto area [41]

$$\Theta^{\text{ext}}_{d,t} = \Theta^{\text{ref}}_{t} + 0.5 \times \varepsilon_{d,t}, \qquad \forall t \in \mathcal{T}$$

with $\varepsilon \sim \mathcal{N}(0,1)$. Numerical parameters for the thermodynamics model are identical among households: $\eta = 1$, $\mu = 0.2$, $C = 3$, and $P^{\text{th, min}} = 0$, $P^{\text{th, max}} = 10$. The inside temperature must be kept in the range $[\Theta^{\min}_d, \Theta^{\max}_d] = [18, 22]$.

The charging of electric vehicles is modelled as a deferrable load with a daily energy requirement of $E^{\min}_d, E^{\max}_d = 10$ as reported in [34]. Charging must happen between 8pm and 5am, and charger limitations are $P^{\min}_d = 1.1$, $P^{\max}_d = 7.7$.

Battery specifications are based on the Tesla powerwall [42]. Energy capacity is $E^{\text{bat, max}}_d = 13.5$, with minimum state of charge $E^{\text{bat, min}}_d = 0$. Charging and discharging limitations are $P^{\text{ch, min}}_d, P^{\text{dis, min}}_d = 0$ and $P^{\text{ch, max}}_d, P^{\text{dis, max}}_d = 5$. Efficiencies are $\eta^{\text{ch}}_d, \eta^{\text{dis}}_d = 0.95$, yielding a 90% round-trip efficiency.

Rooftop solar is modelled as a (negative) curtailable load. The output of a PV system d is given by

$$\tilde{P}^{\text{PV}}_{d,t} = \gamma_r \times \tilde{Q}^{\text{PV}}_{t} \times \zeta_{d,t}, \qquad \forall t \in \mathcal{T}$$

where the normalized output $\tilde{Q}^{\text{PV}}_t$ is Ontario's hourly PV output Q^{PV}_t, divided by its average value over the considered time period. The household-specific scaling factor γ_r is drawn from a uniform distribution $\mathcal{U}(0.5, 1.5)$, and $\zeta \sim \mathcal{U}(0,1)$ is a random noise.

For each household r, the corresponding uncontrollable load d has the following load profile:

$$\tilde{P}_{d,t} = \gamma_r \times \max\left(0, \tilde{Q}^{\text{Ont}}_t + 0.05 \times \varepsilon_{d,t}\right), \qquad \forall t \in \mathcal{T}$$

where $\tilde{Q}^{\text{Ont}}$ is the normalized Ontario hourly provincial load, and $\varepsilon \sim \mathcal{N}(0,1)$ is a white noise.

Finally, we set $P^{\min}_r = 0$, $P^{\max}_r = 10$ for the households' net load constraints. Similarly, the total aggregated load is bounded by $P^{\min}_a = 0$ and $P^{\max}_a = 7.5 \times R$. The price of electricity Π_t is the Hourly Ontario Energy Price (HOEP).

In practice, since neither uncontrollable loads, nor outside temperature, nor PV output, are deterministic, one may use forecasts for these quantities instead. How to obtain such forecasts is beyond the scope of this paper. To handle forecast errors, a receding horizon scheme can be employed, wherein forecasts are updated periodically. When forecasts are updated, e.g. every hour, the aggregation problem is updated accordingly and solved again over the entire time-horizon, e.g. the next 24h. We emphasize that the performance of our algorithm makes such an implementation tractable.

B. Implementation details

Experiments were performed on a 2×Xeon E5-2650V4 2.2Ghz, 256GB RAM computer running Linux. Our implementation was coded in Python 2.7, with CPLEX 12.7 as the linear solver. All CPLEX runs used default parameters and a single thread. Accordingly, an instance is considered solved to optimality when the reported optimality gap is smaller than 10^{-4}, which is the default threshold for CPLEX.

In order to smooth out performance variations, we generated ten different instances for each combination of (R, T, ξ^{dep}). This resulted in a testbed of $(4{\times}7{\times}4){\times}10{=}1120$ instances, and we compare the results obtained by the proposed method (CG), and the centralized formulation (MILP). For the MILP, we used CPLEX with default parameters, a single thread and a time limit of one hour. All instances were found to be feasible. Due to limited computing resources, the column-generation sub-problems were solved serially rather than in parallel. Nevertheless, the duration of each iteration in the distributed setting is given by the RMP computing time, plus the maximum solving time among sub-problems, plus computation overheads. We emphasize that, in practice, each sub-problem would indeed be solved locally by the corresponding household.

Finally, the RMP is initialized by computing, for each resource, a column corresponding to a minimum-peak load schedule. In order to ensure feasibility in the RMP, artificial slack and surplus variables are added to the linking constraints. These artificial variables implement an l_1 penalty with sufficiently large cost, and are thus automatically set to zero once a feasible solution is found. Furthermore, we used a partial pricing strategy that consists in adding at most $0.1 \times R$ columns to the master at each iteration. The 0.1 ratio was found to achieve good performance across a wide range of instances. Similarly, for the recovery heuristic, a random 10% of the resources are selected at each step.

C. Performance analysis

Performance statistics are presented for MILP and CG in Tables II and III respectively. Since both methods behaved consistently across the different values of ξ^{dep}, we only report

Table II
MILP STATISTICS FOR $T = 24$, $\xi^{\text{DEP}} = 0.66$

R	Variables (M) Bin.	Cont.	Solved (%)	Time (s) Root	Total	B&B nodes	Root % gap
1024	0.13	0.19	100	9.6	40.1	1.5	0.00
1536	0.19	0.29	100	20.6	132.9	11.5	0.00
2048	0.26	0.39	100	30.2	291.0	108.1	0.00
3072	0.39	0.57	100	65.4	649.8	394.4	0.00
4096	0.52	0.77	100	147.3	1263.7	562.8	0.00
6144	0.78	1.15	70	314.8	2907.1	550.1	0.00
8192	1.04	1.54	60	534.6	2937.4	664.3	0.00

Table III
CG STATISTICS FOR $T = 24$, $\xi^{\text{DEP}} = 0.66$

R	Col. (k)	Iter.	Solved (%)	Time (s) Master	Pricing	Total	Root % gap
1024	5.3	45.7	100	0.6	1.7	3.8	0.00
1536	8.0	46.0	100	1.1	1.7	5.1	0.00
2048	10.6	46.1	100	1.8	1.8	6.8	0.00
3072	16.2	46.6	100	3.8	1.9	10.6	0.00
4096	21.6	47.2	100	6.4	2.0	15.0	0.00
6144	31.7	46.3	100	13.5	2.0	25.2	0.00
8192	43.4	47.6	100	25.1	2.1	40.6	0.00

results for the case $\xi^{\text{dep}} = 0.66$. For the MILP formulation, Table II reports the number of binary and continuous variables (Bin. and Cont. respectively, in millions) of the problem, the proportion of instances solved to proven optimality, total and root computing times (in seconds), the number of nodes in the branch-and-bound tree, and the root gap (in %). For the CG method, Table III reports the number of columns generated (Col., in thousands), the total number of CG iterations to reach optimality (Iter.), the proportion of instances solved to proven optimality, computing times (in seconds), and the root gap (in %). Root gaps are given by

$$\text{gap} = \frac{|z^* - \underline{z}|}{|z^*|},$$

where z^* is the value of the best known integer solution found by either algorithm, and $\underline{z}$ is the value of the root relaxation. Finally, all reported averages are geometric means.

As expected, CPLEX solved the MILP formulation for smaller instances, but systematically reached the time limit for the larger ones. More specifically, CPLEX failed to find a feasible solution for 162 instances, roughly corresponding to the instances with more than two million binary variables. Note that the quality of the lower bound for MILP is not an issue here. On the opposite, the solver's capability of exploring the branch-and-bound nodes in a reasonable amount of time, in order to provide good feasible solutions, is affected by the problem's size. In comparison, CG was able to solve all 1120 instances to proven optimality. Furthermore, computing times for both methods are displayed in Figure 1. For all values of T, CG exhibits a more scalable behaviour than MILP, and achieves speed-ups of up to two orders of magnitude. These results confirm that, even if it may be tractable for small-size systems, a centralized approach fails to handle large numbers of resources.

To further analyse the scalability of CG, Figure 2 shows the number of CG iterations to reach convergence, corresponding to the number of times the RMP is solved. Here, the number of iterations is more relevant than raw computation times, since the latter depends on machine specifications and on the solver's performance. On one hand, an increase in the length of the time-horizon T leads, as expected, to an increase in the number of iterations. On the other hand, as was hinted at in Table III, the number of iterations appears to be independent

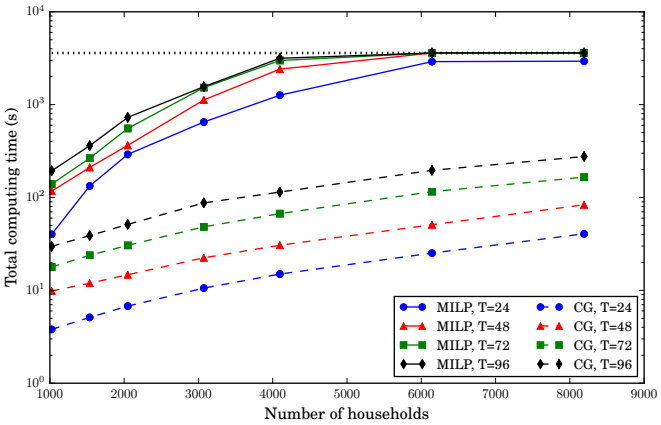

Figure 1. Average computing times for CG and MILP formulations, with $\xi^{\text{dep}} = 0.66$. The horizontal dotted line depicts the one-hour time limit for MILP.

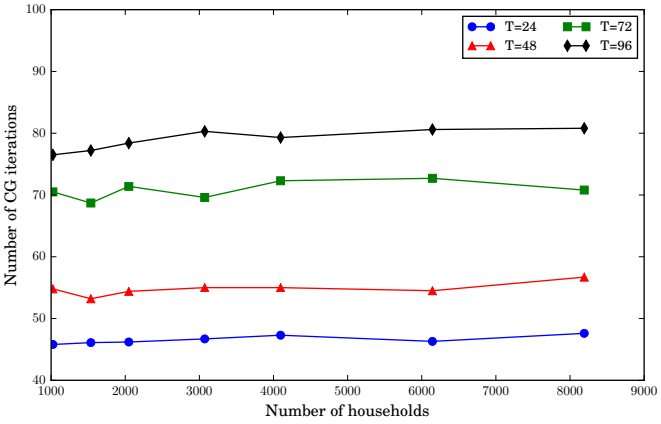

Figure 2. Average number of CG iterations, including the diving heuristic. $\xi^{\text{dep}} = 0.66$.

of the number of households. This remarkable behaviour is due to similarities between sub-problems. This explanation is corroborated by the fact that larger values of ξ^{dep}, which only affects the distribution of households, and thus of sub-problems, resulted in more iterations. Overall, CG is sensitive to the distribution of resources, rather than their number.

Finally, the proportions of time spent solving the master problem and sub-problems, respectively, are displayed in Fig-

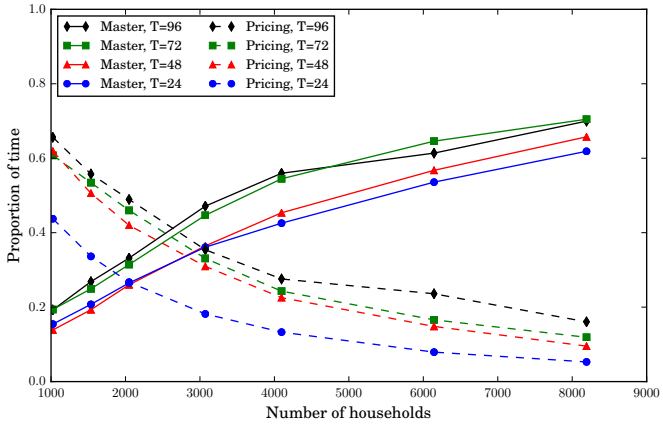

Figure 3. Time spent solving the master problem and pricing sub-problems, as a fraction of total computing time. $\xi^{\text{dep}} = 0.66$.

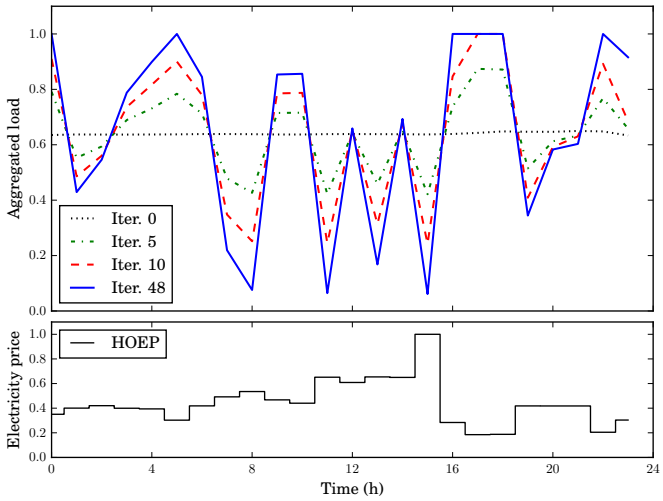

Figure 4. Normalized Hourly Ontario Electricity Price (HOEP, lower graph) and normalized aggregated load (upper graph) over the time horizon. The aggregated load is displayed at iterations $0, 5, 10$ and 48 (last iteration) of the CG algorithm. $R = 1024, T = 24, \xi^{dep} = 0.66$.

ure 3. Computation overheads can be significant, but they are highly dependent on the implementation. Therefore, we factor them out of the plots. Clearly, as the number of resources increases, most of the time is spent solving the RMP. Indeed, since sub-problems are solved in parallel, the number of resources has little influence on the duration of the pricing step. This is consistent with computing times reported in Table III. Conversely, the size of the RMP is given by the number of columns generated. Therefore, as R increases, so does the size of the RMP and the associated computational cost. Consequently, solving the RMP is a computation bottleneck for CG, and currently constitutes the main limitation to the algorithm's scalability.

D. Evolution of the aggregated load

The evolution of the aggregated load is displayed in Figure 4, together with the market price of electricity over the considered time horizon. The normalized load takes value 1 (resp. 0) when the aggregated load reaches its upper bound $P_a^{\max}$ (resp. lower bound $P_a^{\min}$).

The initial load profile (Iter. 0) corresponds to each resource computing its minimum peak load schedule. Since the objective is to minimize the cost of purchasing electricity from the grid, we expect consumption to be shifted to periods when electricity is cheapest, while periods of high price should result in low demand. The evolution of the load depicted in Figure 4 shows that this is indeed the case. At the optimum (Iter. 46), consumption is highest at periods $t = 0, 5, 16, 17, 18$ and 22, which correspond to periods of lower prices.

Finally, the upper bound on the aggregated load is reached, e.g. for $t = 16, 17, 18$ at the optimum, but never violated. This demonstrates that the algorithm is effectively enforcing the linking constraints during the optimization process. This last feature is desirable, since a feasible solution will be available if the algorithm is stopped prematurely, e.g. due to time restrictions.

E. Further discussion

As shown in Tables II and III, root gaps for CG and MILP were often found to be lower than 10^{-4}, meaning both formulations have essentially zero integrality gap. This observation carries over to the rest of the dataset: for MILP, root gaps were always less than 0.02%, and smaller than 0.01% in 843 instances out of 1120. For CG, the largest recorded root gap was 0.008%. Moreover, root gaps tended to be smaller as the number of resources increased.

We now assess whether low integrality gaps reflect intrinsic properties of the problem at hand, or arise from our numerical data. To that end, we increased the batteries' minimum power ratios $P^{\text{ch, min}}/P^{\text{ch, max}}$ (resp. $P^{\text{dis, min}}/P^{\text{dis, max}}$) from 0, as in our initial tests, to 90%. This is done by setting the value of $P^{\text{ch, min}}$ (resp. $P^{\text{dis, min}}$) accordingly. Figure 5 displays the resulting evolution of computing time (left axis) and integrality gap (right axis). Results are reported for $R = 1024$, $T = 24$ and $\xi^{\text{dep}} = 0.66$, and similar behaviour were observed for other settings. Figure 5 shows that the MILP integrality gaps increases considerably, from 0.005% to over 2%. This resulted in an increase in the number of branching nodes, which we do not report for lack of space, and in computing time. On the other hand, although the CG gap increased, it remained below 0.01%. Moreover, the increase in computing time for CG is caused only by longer solving times for sub-problems. The number of CG iterations did not increase, nor did solving time for the RMP. Overall, CG appears to be more robust than MILP. This robustness is explained by the fact that changes in the resources' operation only affect sub-problems for CG while, for MILP, the entire problem structure may be affected.

V. Conclusion

In this paper, we have considered the problem of coordinating the operation of multiple DERs, and focused on the challenges raised by the presence of discrete decisions in the resources' operation, such as on-off constraints. We showed that this problem can be formulated as a centralized MILP, which is however intractable for large systems.

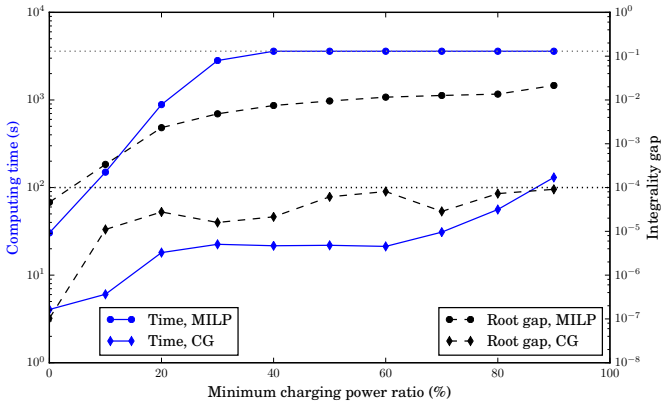

Figure 5. Average computing time (blue, left-hand axis scale) and root gap (black, right-hand axis scale) with respect to minimum charging power ratio. Dotted horizontal lines denote the one-hour time limit for MILP (top, in gray), and the 10^{-4} optionality threshold for integrality gaps (bottom, in black). Results obtained with $R = 1024$, $T = 24$ and $\xi^{\text{dep}} = 0.66$.

We have proposed an exact methodology for solving this MILP efficiently, based on Dantzig-Wolfe decomposition and Column Generation. In particular, we have shown that the proposed formulation is technology agnostic and can be implemented in a decentralized fashion, thus yielding high scalability while providing enhanced privacy to the resources. We have also reported on extensive computational results, which demonstrate the efficiency and robustness of our approach.

Future work will focus on integrating demand and price uncertainty in the formulation, as well as developing acceleration strategies to improve practical performance.

ACKNOWLEDGEMENTS

This work was supported by the NSERC Energy Storage Technologies Network (NEST Net). The authors are grateful to the anonymous reviewers for valuable feedback that helped improve this paper.

REFERENCES

[1] M. H. Albadi and E. F. El-Saadany, "A summary of demand response in electricity markets," *Electric Power Systems Research*, vol. 78, no. 11, pp. 1989–1996, 2008.

[2] IESO, "Distributed energy resources," accessed: 03-09-2018. [Online]. Available: http://www.ieso.ca/en/learn/ontario-power-system/a-smarter-grid/distributed-energy-resources

[3] L. Gkatzikis, I. Koutsopoulos, and T. Salonidis, "The role of aggregators in smart grid demand response markets," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1247–1257, 2013.

[4] P. Siano, "Demand response and smart grids - a survey," *Renewable and Sustainable Energy Reviews*, vol. 30, no. Supplement C, pp. 461 – 478, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032113007211

[5] R. Deng, Z. Yang, M. Y. Chow, and J. Chen, "A survey on demand response in smart grids: Mathematical models and approaches," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 570–582, June 2015.

[6] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, "A survey on demand response programs in smart grids: Pricing methods and optimization algorithms," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 152–178, 2015.

[7] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen, "A bilevel model for electricity retailers' participation in a demand response market environment," *Energy Economics*, vol. 36, pp. 182 – 197, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0140988312003477

[8] P. Yang, G. Tang, and A. Nehorai, "A game-theoretic approach for optimal time-of-use electricity pricing," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 884–892, May 2013.

[9] D. T. Nguyen, H. T. Nguyen, and L. B. Le, "Dynamic pricing design for demand response integration in power distribution networks," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3457–3472, Sept 2016. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7372487

[10] E. Alekseeva, L. Brotcorne, S. LEPAUL, and A. MONTMEAT, "bilevel approach to optimize electricity prices," *Yugoslav Journal of Operations Research*, 2018. [Online]. Available: http://www.doiserbia.nb.rs/img/doi/0354-0243/2018%20OnLine-First/0354-02431800002A.pdf

[11] A. H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Transactions on Smart Grid*, vol. 1, no. 3, pp. 320–331, Dec 2010.

[12] Z. Baharlouei, H. Narimani, and H. Mohsenian-Rad, "Tackling co-existence and fairness challenges in autonomous demand side management," in *2012 IEEE Global Communications Conference (GLOBECOM)*, Dec 2012, pp. 3159–3164.

[13] L. Chen, N. Li, S. H. Low, and J. C. Doyle, "Two market models for demand response in power networks," in *2010 First IEEE International Conference on Smart Grid Communications*, Oct 2010, pp. 397–402.

[14] N. Li, L. Chen, and S. H. Low, "Optimal demand response based on utility maximization in power networks," in *2011 IEEE Power and Energy Society General Meeting*, July 2011, pp. 1–8.

[15] H. Zhong, L. Xie, and Q. Xia, "Coupon incentive-based demand response: Theory and case study," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1266–1276, May 2013.

[16] J. Li, Z. Wu, S. Zhou, H. Fu, and X.-P. Zhang, "Aggregator service for pv and battery energy storage systems of residential building," *CSEE Journal of Power and Energy Systems*, vol. 1, no. 4, pp. 3–11, 2015.

[17] M. Parvania, M. Fotuhi-Firuzabad, and M. Shahidehpour, "Optimal demand response aggregation in wholesale electricity markets," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 1957–1965, 2013.

[18] ——, "Iso's optimal strategies for scheduling the hourly demand response in day-ahead markets," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2636–2645, 2014.

[19] Z. Zhu, J. Tang, S. Lambotharan, W. H. Chin, and Z. Fan, "An integer linear programming based optimization for home demand-side management in smart grid," in *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, Jan 2012, pp. 1–5.

[20] P. Chavali, P. Yang, and A. Nehorai, "A distributed algorithm of appliance scheduling for home energy management system," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 282–290, Jan 2014.

[21] T. Logenthiran, D. Srinivasan, and T. Z. Shun, "Demand side management in smart grid using heuristic optimization," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1244–1252, Sept 2012.

[22] K. Paridari, A. Parisio, H. Sandberg, and K. H. Johansson, "Demand response for aggregated residential consumers with energy storage sharing," in *2015 54th IEEE Conference on Decision and Control (CDC)*, Dec 2015, pp. 2024–2030.

[23] M. A. A. Pedrasa, T. D. Spooner, and I. F. MacGill, "Coordinated scheduling of residential distributed energy resources to optimize smart home energy services," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 134–143, Sept 2010.

[24] M. Kraning, E. Chu, J. Lavaei, S. Boyd *et al.*, "Dynamic network energy management via proximal message passing," *Foundations and Trends® in Optimization*, vol. 1, no. 2, pp. 73–126, 2014.

[25] J. Rivera, P. Wolfrum, S. Hirche, C. Goebel, and H. A. Jacobsen, "Alternating direction method of multipliers for decentralized electric vehicle charging control," in *52nd IEEE Conference on Decision and Control*, Dec 2013, pp. 6960–6965.

[26] N. Gatsis and G. B. Giannakis, "Residential demand response with interruptible tasks: Duality and algorithms," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, Dec 2011, pp. 1–6.

[27] ——, "Decomposition algorithms for market clearing with large-scale demand response," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 1976–1987, Dec 2013.

[28] S. J. Kim and G. B. Giannakis, "Scalable and robust demand response with mixed-integer constraints," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2089–2099, Dec 2013.

[29] S. Mhanna, A. C. Chapman, and G. Verbic, "A fast distributed algorithm for large-scale demand response aggregation," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 2094–2107, July 2016.

[30] P. M. Namara and S. McLoone, "Hierarchical demand response using dantzig-wolfe decomposition," in *IEEE PES ISGT Europe 2013*, Oct 2013, pp. 1–5.
[31] H. A. Toersche, A. Molderink, J. L. Hurink, and G. J. M. Smit, "Column generation based planning in smart grids using triana," in *IEEE PES ISGT Europe 2013*, Oct 2013, pp. 1–5.
[32] M. H. Amini, P. McNamara, P. Weng, O. Karabasoglu, and Y. Xu, "Hierarchical electric vehicle charging aggregator strategy using dantzig-wolfe decomposition," *IEEE Design Test*, vol. PP, no. 99, pp. 1–1, 2017.
[33] H. Harb, J.-N. Paprott, P. Matthes, T. Schütz, R. Streblow, and D. Müller, "Decentralized scheduling strategy of heating systems for balancing the residual load," *Building and Environment*, vol. 86, pp. 132 – 140, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360132314004260
[34] M. Beaudin and H. Zareipour, "Home energy management systems: A review of modelling and complexity," *Renewable and Sustainable Energy Reviews*, vol. 45, pp. 318 – 335, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032115000568
[35] G. T. Costanzo, G. Zhu, M. F. Anjos, and G. Savard, "A system architecture for autonomous demand side load management in smart buildings," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 2157–2165, Dec 2012.
[36] J. P. Holdman, *Heat Transfer*, 10th ed. McGraw-HIll, 1997.
[37] J. A. Gomez and M. F. Anjos, "Collaborative demand-response planner for smart buildings," GERAD, Tech. Rep. Technical Report G-2017-39, 2017. [Online]. Available: https://www.gerad.ca/en/papers/G-2017-15/view
[38] G. B. Dantzig and P. Wolfe, "Decomposition principle for linear programs," *Operations Research*, vol. 8, no. 1, pp. 101–111, 1960. [Online]. Available: https://doi.org/10.1287/opre.8.1.101
[39] G. Desaulniers, J. Desrosiers, and M. M. Solomon, *Column generation*, 1st ed., ser. GERAD 25th anniversary. Springer Science & Business Media, 2006, vol. 5.
[40] IESO, "Power data." [Online]. Available: http://www.ieso.ca/power-data
[41] "Historical data," toronto City station, ID 6158355. [Online]. Available: http://climate.weather.gc.ca/historical_data/search_historic_data_e.html
[42] Tesla, "Powerwall specifications," accessed: 08-10-2017. [Online]. Available: https://www.tesla.com/support/powerwall

Miguel F. Anjos (M 2007, SM 2018) is Professor at Polytechnique Montreal where he holds the NSERC-Hydro-Quebec-Schneider Electric Industrial Research Chair, and an Inria International Chair. He is a Licensed Professional Engineer in Ontario, Canada, and earned his degrees at McGill University, Stanford University, and the University of Waterloo. He served as Editor-in-Chief of Optimization and Engineering, and serves on several editorial boards. His allocades include a Canada Research Chair, the Méritas Teaching Award, a Humboldt Research Fellowship, the title of EUROPT Fellow, and the Queen Elizabeth II Diamond Jubilee Medal. He is a Fellow of the Canadian Academy of Engineering.

Andrea Lodi has been full professor of Operations Research at DEI, University of Bologna between 2007 and 2015. Since 2015 he is Canada Excellence Research Chair in "Data Science for Real-time Decision Making" at Polytechnique Montréal.

His main research interests are in Mixed-Integer Linear and Nonlinear Programming and Data Science and his work has received several recognitions including the IBM and Google faculty awards. He has been network coordinator and principal investigator of two large EU projects/networks, and, since 2006, consultant of the IBM CPLEX research and development team. He is the co-principal investigator of the project "Data Serving Canadians: Deep Learning and Optimization for the Knowledge Revolution" and scientific co-director of IVADO, the Montréal Institute for Data Valorization.

Mathieu Tanneau received the M.Sc from École polytechnique, France, in 2017, and is currently pursuing the Ph.D. degree at Polytechnique Montréal, Canada.

His research interests are in distributed methods for large scale mixed-integer optimization, as well as the integration of machine learning techniques within optimization frameworks.